RESEARCH PAPER

Artificial DNA: PNA & XNA 2:1, 33-37; January/February/March 2011; © 2011 Landes Bioscience

A novel pseudo-complementary PNA G-C base pair

Anne G. Olsen,[1,2] Otto Dahl,[2] Asger B. Petersen,[1,3] John Nielsen[3] and Peter E. Nielsen[1,4,*]

[1]Department of Cellular and Molecular Medicine; Faculty of Health Sciences; The Panum Institute; [2]Department of Chemistry; [3]Faculty of Life Sciences; Department of Basic Sciences and Environment; [4]Faculty of Pharmaceutical Sciences; Department of Medicinal Chemistry; University of Copenhagen; Copenhagen, Denmark

Key words: DNA recognition, hybridization, nucleobases, synthesis, PNA

Abbreviations: DMF, *N,N*-dimethylformamide; DMSO, dimethylsulfoxide; ES, electrospray; FAB, fast atom bombardment; Fmoc, fluorenylmethoxycarbonyl; MALDI, matrix assisted laser desorption ionisation; PNA, peptide nucleic acids; TOF, time of flight

Pseudo-complementary oligonucleotide analogues and mimics provide novel opportunities for targeting duplex structures in RNA and DNA. Previously, a pseudo-complementary A-T base pair has been introduced. Towards sequence unrestricted targeting, a pseudo-complementary G-C base pair consisting of the unnatural nucleobases *N*6-methoxy-2,6-diaminopurine (previously described in a DNA context) and *N*4-benzoylcytosine is now presented for design of pseudo-complementary PNA oligomers (pcPNAs).

Introduction

A pair of pseudo-complementary oligonucleotides are formally sequence complementary, but have significantly reduced affinity for forming duplexes with each other due to chemical modification. However, they retain high affinity for complementary natural DNA or RNA targets. Such pseudo-complementary oligonucleotides are of major interest for sequence specific targeting of duplex DNA by double duplex invasion strategies,[1-10] and as efficient sequence palindromic (self-complementary) hybridization probes for hairpin forming in RNA targets.[11-13]

Previous work has shown that PNA or DNA oligomers with pseudo-complementary properties are obtained upon substitution of A-T bases by 2,6-diaminopurine (D)-thiouracil (sU) (or thiothymine) bases, whereas analogous pseudo-complementary G-C base pairs have proven elusive.[13]

In order to approach an effective pseudo complementary G-C base pair we considered that the unnatural nucleobases *N*6-methoxy-2,6-diaminopurine (previously described in a DNA context),[14] and *N*4-benzoylcytosine[15] could indeed constitute such a pair. These two modified nucleobases recognize cytosine and guanine, respectively, but a steric clash between the methoxy group of *N*6-methoxy-2,6-diaminopurine (K*) and the benzoyl group *N*4-benzoylcytosine (C^{Bz}) would be anticipated (**Fig. 1**). Thus we have synthesized a K* PNA monomer, incorporated this in PNA oligomers and studied the pseudo-complementary properties of the K*-C^{Bz} base pair by PNA-PNA and PNA-DNA hybridization analyses.

Results and Discussion

The synthesis of the monomer containing the K* nucleobase is shown in **Scheme 1**. Compound 1 was synthesised from 2-amino-6-chloropurine according to a previously published procedure.[16] Subsequently, 1 was condensed with *tert*-butyl *N*-[2-(*N*-9-fluorenylmethoxycarbonyl)-aminoethyl] glycinate[17] using DCC and DhbtOH as coupling reagents to obtain 2 in 70% yield. The *tert*-butyl ester functionality on 2 was cleaved using trifluoroacetic acid at 0°C to afford the free acid 3 in 63% yield. Conversion of 4 into the *N*6-methoxy monomer 4 was performed with methoxyamine[18] in dioxane at 90°C to obtain the title compound in 57% yield. The monomer 4 was incorporated into PNA oligomers under solid phase synthesis conditions using Fmoc chemistry.[17]

The nucleobase recognition properties of K* was studied by incorporating the monomer 4 in place of G into a PNA oligomer. The thermal stability of duplexes of this PNA as compared to the guanine containing PNA and sequence complementary DNA in which the position opposite the K* was varied, was analysed by thermal denaturation (**Table 1**). It is clearly seen that while guanine of course very strongly prefers to pair with cytosine, the K* base pairs equally well with cytosine and thymine, and very poorly with adenine or guanine (ΔTm ~ 20°C). It is also observed that the K* containing PNA-DNA duplex is somewhat less stable (ΔTm ~ 10°C) than the unmodified duplex. This decrease is in line with the results obtained by Brown et al. in a DNA-DNA context (ΔTm/mod. = -9.0°C) and may be ascribed to the necessary change in conformation of the methoxyamino group from

*Correspondence to: Peter E. Nielsen; Email: ptrn@sund.ku.dk
Submitted: 02/01/11; Revised: 03/18/11; Accepted: 03/22/11
DOI: 10.4161/adna.2.1.15554

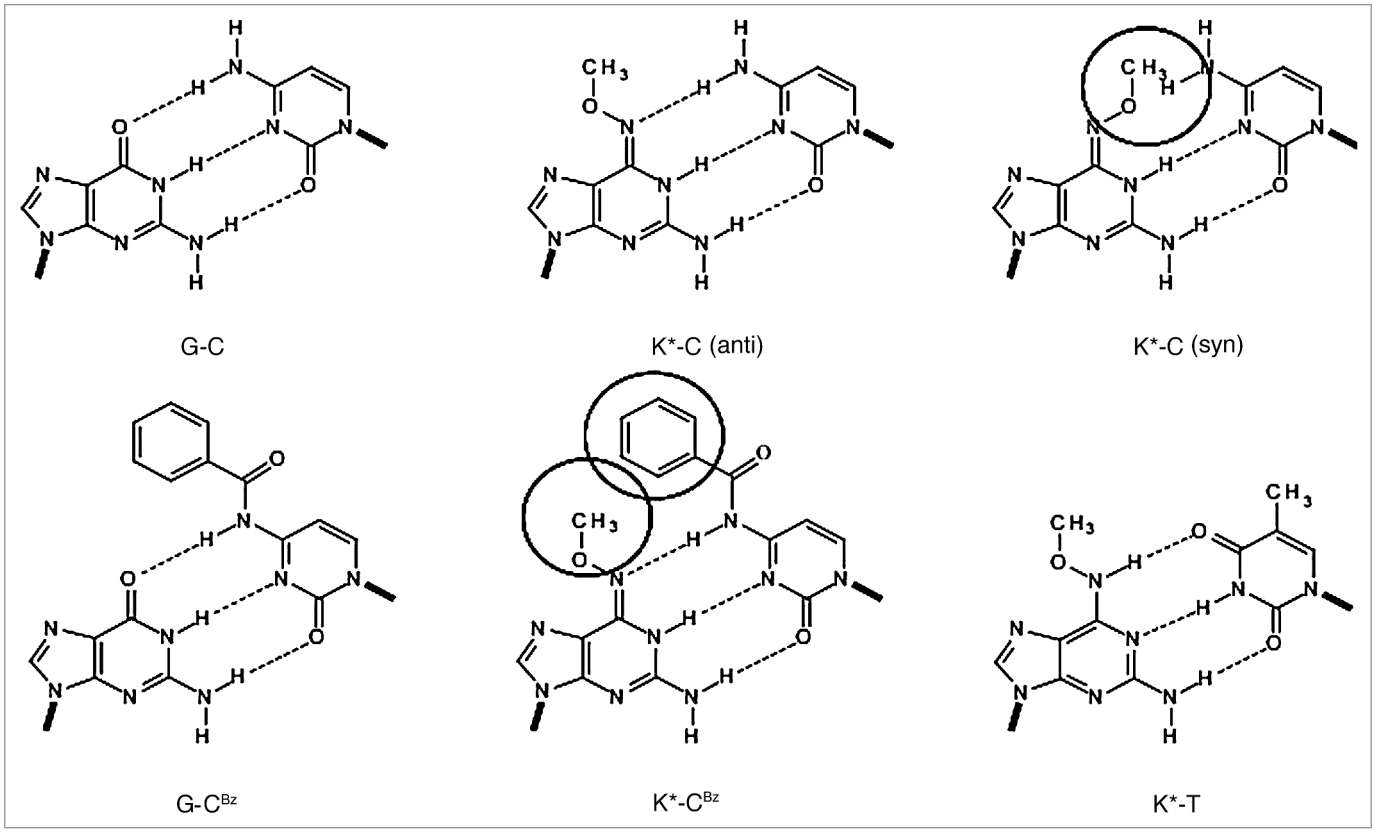
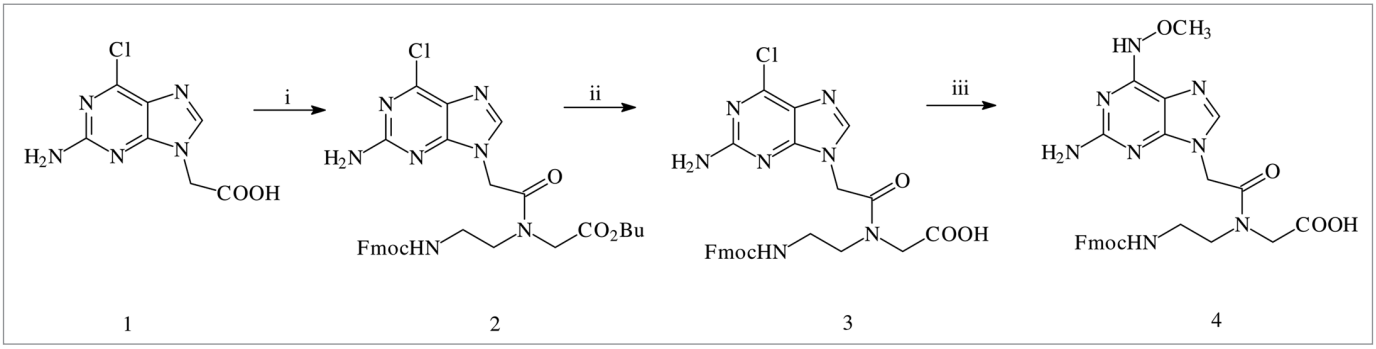


Figure 1. Schematic drawing of G-C, K*-C, G-C^{Bz} and K*-C^{Bz} base pairs showing how a steric clash, that can potentially lead to double duplex invasion, is expected between K* and C^{Bz}. The K* base is shown in the imino tautomeric form and the methoxy group adopts the anti conformation. To the right (upper corner) the K*-C base pair is shown with the methoxy group adopts the syn conformation preventing H-bond formation and creating a steric clash. The K*-T base pair with the K* base i in the amino tautomeric form and the methoxy group adopting the anti conformation is also shown in the right/down corner.



Scheme 1. (i) FmocNHCH$_2$CH$_2$-NHCH$_2$CO$_2$^{t}Bu, DCC, DhbtOH, DMF, 0°C then rt (70%); (ii) TFA, 0°C (63%); (iii) MeONH$_2$•HCl, DIPEA, 1,4-dioxane, 90°C (57%). The K* base in **4** is depicted in the amino tautomeric form showing pairing with T in a Watson-Crick manner, the methoxy group adopts the anti conformation.

syn to anti in order to bind cytosine in a Watson-Crick manner (**Fig. 1**). Likewise the promiscuous T/C recognition has been observed in a DNA context,[14] and can be explained by the ability of K* to exist in two tautomeric forms, capable of binding cytosine (**Fig. 1**) and thymine (**Scheme 1**), respectively.

The critical question was whether the K* base would more strongly than guanine discriminate cytosine versus the sterically challenging *N*4-benzoylcytosine (C^{Bz}) (**Fig. 1**), and thereby provide a candidate for a G-C pseudo-complementary base pair, i.e., the C^{Bz}-K*pair. This was studied by measuring the stabilities of the duplexes between the PNA and DNA oligomers and a C^{Bz} substituted PNA oligomer (**Fig. 2 and Table 2**). Clearly the stability of a K* substituted PNA duplex is affected significantly more by C^{Bz} substitution at the opposing base position as compared to the normal C containing PNA. The Tm is reduced by 12.8°C in case of K* containing PNA, as compared to only 2.7°C

for the normal PNA. These results clearly support the notion of a steric clash between the methoxy and the benzoyl groups (**Fig. 1**), and thus support the possibility of an effective C^{Bz}-K*pseudo-complementary pair.

However, we also note that the K* base per se does compromise the overall stability of the PNA-DNA duplex somewhat more than it affects PNA-PNA duplex stability (ΔTm = -11.6°C versus -7°C). Nonetheless. the PNA-PNA destabilization by K*-C^{Bz} base combination is still more pronounced (ΔTm = -19.8°C compared to a normal G-C base pair) (**Table 2**), even considering the slight destabilization of the PNA-DNA duplex by the C^{Bz} base (ΔTm = -2.1°C). Ideally, the nucleobases of a GC pseudo complementary base pair in a PNA oligomer should leave the DNA sequence recognition properties of this PNA oligomer unchanged, while they should significantly destabilize the PNA-PNA duplex. Therefore incorporation of more than one or a few K* bases in a PNA oligomer might not be sufficiently effective and the K*C^{Bz} base pair consequently is not an ideal pseudo-complementary base pair for very GC rich sequences. Thus further experiments are warranted examining the combined effect of K*C^{Bz} base pairs together with DsU base pairs in pseudo complementary PNA oligomers (pcPNAs) of lower than 50% TA content.

The present results clearly show that steric hindrance between properly *N-(O)* substituted diaminopurine and *N*-substituted cytosine may indeed be exploited for discovery of pseudo-complementary GC base pairs, and the data obtained with he K*-C^{Bz} pair may be used for future development of an optimal pair, e.g., by exchanging the *O*-methyl group with various other (alkyl) substituents, and/or adding substituents to the benzoyl group or exploiting other (aromatic) acids.

Table 1. Thermal stability (T_m) of PNA-DNA duplexes

DNA sequence	PNA	Tm (°C)	ΔTm/mod. (°C)[a]
5'-dAGTG**C**TCTAC-3'	X = **G**	61.9	
5'-dAGTG**T**TCTAC-3'	X = **G**	40.2	-21.7
5'-dAGTG**A**TCTAC-3'	X = **G**	34.2	-27.7
5'-dAGTG**G**TCTAC-3'	X = **G**	36.3	-25.6
5'-dAGTG**C**TCTAC-3'	X = **K***	50.3	
5'-dAGTG**T**TCTAC-3'	X = **K***	51.4	+1.1
5'-dAGTG**A**TCTAC-3'	X = **K***	30.1	-20.2
5'-dAGTG**G**TCTAC-3'	X = **K***	29.2	-21.1

[a]ΔTm is relative to hybridization data for the matched PNA-DNA duplexes. PNA sequence:H-GTAGA**X**CACT-LysNH$_2$.

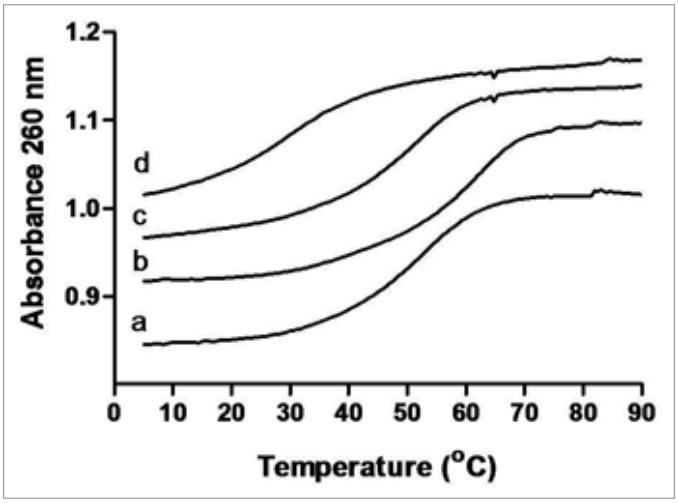


Figure 2. Thermal denaturation profiles (Tm, melting curves) of PNA H-GTAGA**K***CACT-LysNH$_2$ duplexes with (a) H-AGTG**C^{Bz}**TCTAC-LysNH$_2$, (b) PNA H-AGTG**C**TCTAC-LysNH$_2$, (c) DNA 5'-dAGTG**T**TCTAC-3'and (d) DNA 5'-dAGTG**G**TCTAC-3'.

Materials and Methods

For reactions carried out under argon, the glassware was dried before use, employing the following procedure: Heating was applied with a heat gun while the glassware was evacuated (16 torr), the evacuation was followed by argon flushing. Evacuation and flushing were applied several times under continuous heating, and finally the glassware was allowed to cool to room temperature under argon before use. All reagents were purchased from major suppliers and used without further purification unless otherwise noted. All solvents were analytical grade and dried when necessary over either 3 Å or 4 Å activated molecular sieves. Water content was determined to be less than 20 ppm by Karl Fischer titration on a Metrohm 831 KF-Coulometer. Thin Layer Chromatography (TLC) was carried out on commercially available pre-coated plates (Merck 60 F_{254}). Dry Column Vacuum Chromatography (DCVC) was performed according to published procedure.[19] Reverse phase DCVC was performed using the same procedure, with the modifications that Waters RP 18 (0.055–0.105 mm) silica gel was employed, and the compound was applied on the column dissolved in a small amount 60:40, water:acetonitrile. ^{1}H and ^{13}C NMR spectra were recorded on a Varian 400 MHz or a Bruker Avance 300 Fourier transform spectrometer using an internal deuterium lock. Solvents were used as internal standards when assigning NMR spectra (δ_H: DMSO-d_6 2.50; δ_C: DMSO-d_6 39.4 ppm). J values are rounded to the nearest 0.5 Hz, minor rotamers are given in brackets. Purity determined by ^{1}H NMR, solvent residues not included, all yields are corrected. FAB and TOF-HRMS-ES$^+$ mass spectra were recorded in positive ion mode on a Jeol JMS-HX/HX110A and a Micromass Q-Tof instrument, respectively, and ESI$^+$ mass spectra were recorded on a Micromass LCT apparatus with an AP-ESI probe and either Leu-enkephalin (556.2771 g/mol) or Fmoc-His(Trt)-OH (620.2549 g/mol) as reference. Preparative HPLC was run on a Waters preparative HPLC system consisting of a 2525 Binary Gradient Module, an UV Fraction Manager, a 2996 Photo Diode Array, and a 2767 Sample Manager. The column was a Waters XTerra Prep MS C_{18} (5 μm) ODB Column 19 x 100 mm. The flow was 15 mL/min. with a linear water:acetonitrile gradient (both solvents containing 0.1% TFA).

(2) **{[2-(2-Amino-6-chloro-purin-9-yl)-acetyl]-[2-(9H-fluoren-9-ylmethoxycarbonyl-amino)-ethyl]-amino}-acetic acid tert-butyl ester.** To a stirred solution of 1 (0.50 g, 2.2 mmol) and DhbtOH (0.36 g, 2.2 mmol) in DMF (10 cm3) at 0°C under

Table 2. Thermal stability data associated with combinations of DNA and PNA strands relevant for revealing the double duplex invasion potential

10-mer PNA/DNA sequence	DNA		PNA X = C		PNA X = C^{Bz}	
	Tm (°C)	ΔTm/mod. (°C)[a]	Tm (°C)	ΔTm/mod. (°C)[a]	Tm (°C)	ΔTm/mod. (°C)[a]
H-GTAGA**G**CACT-LysNH_2	61.9		72.0		69.3	
H-GTAGA**K***CACT-LysNH_2	50.3	-11.6	65.0	-7.0	52.2	-17.1
5'-dGTAGA**G**CACT-3'	41.1		50.1		48.0	

DNA complement: 5'-dAGTG**C**TCTAC-3', PNA complement: H-AGTG**X**TCTAC-LysNH_2. [b]ΔTm is relative to hybridization data for the guanine containing control.

argon, was added FmocNHCH_2CH_2NHCH_2CO_2^{t}Bu (0.87 g, 2.0 mmol) [the hydrochloride was dissolved in DCM (50 cm^3) extracted with saturated aqueous $NaHCO_3$ (2 x 20 cm^3), dried with $MgSO_4$ and evaporated in vacuo before transferring to the reaction mixture with DMF (5 cm^3)] and DCC (0.980 g, 4.8 mmol). The mixture was allowed to reach room temperature over night. The precipitate was separated from the organic phase and extracted with DMF (5 cm^3) and DCM (40 cm^3). The combined organic phase was extracted with saturated aqueous $NaHCO_3$ (2 x 30 cm^3), saturated aqueous $KHSO_4$ (30 cm^3), brine (20 cm^3), dried with Na_2SO_4 and evaporated in vacuo; DCVC (ø 2 cm x 4 cm) 100, 60, 30 and 0% heptane in ethyl acetate 20 cm^3 fractions followed by 100% ethyl acetate→20% ethyl acetate in 80% methanol 5% increments/20 cm^3 fraction yielded 0.86 g (71%) light yellow solid (contained 10 mol% DCU). R_f 0.35 (ethyl acetate); δ_H (300 MHz, DMSO-*d6*) 7.97 (1H, s, *H*8), 7.89 (2H, d, *J* 7.5 Hz, fluorenyl-*H*4 & 5), 7.68 (7.66) (2H, d, *J* 7 (6.5) Hz, fluorenyl-*H*1 & 8), 7.49-7.25 (5H, m, fluorenyl-*H*2,3,6,7 & Fmoc-N*H*), 6.89 (2H, br s, NH_2), 5.11 (4.93) (2H, s, CH_2N9), 4.42-4.17 (3.96) (5H, m, C*H*CH_2 & CH_2COOH), 3.55-3.01 (4H, m, CH_2CH_2), (1.47) 1.36 (9H, s, CH_3). *m*/*z* (FAB$^+$-MS) 606 (M^+).

(3) **{[2-(2-Amino-6-chloro-purin-9-yl)-acetyl]-[2-(9*H*-fluoren-9-ylmethoxycarbonyl-amino)-ethyl]-amino}-acetic acid.** Compound 2 (1.48 g, 2.44 mmol) was stirred in trifluoroacetic acid (45 ml) at 0°C for 6½ h. The solution was evaporated to dryness in vacuo and co-evaporated once with toluene (30 ml). The residue was dissolved in acetone and the crude product was purified by DCVC using *n*-heptane-ethyl acetate (100:0 to 0:100 v/v) followed by ethyl acetate-methanol (100:0 to 70:30 v/v) as eluent to give 3 (0.846 g, 63%) as a white solid. δ_H (400 MHz, DMSO-*d6*) 12.8 (1H, br s, COO*H*), 7.97 (1H, s, *H*-8), 7.87 (2H, d, *J* 7.5 Hz, Fmoc), 7.67 (2H, t, *J* 7.5 Hz, Fmoc), 7.47-7.25 (5H, m, Fmoc, FMocN*H*), 6.88 (2H, br s, NH_2), 5.12 (4.95) (2H, s, CH_2N9), 4.38-4.18 (3H, m, Fmoc), 4.01 (2H, s, COCH_2N), 3.53-3.10 (4H, m, CH_2CH_2); δ_C (100.6 MHz, DMSO-*d6*) 170.8, 167.0, 166.5, 159.8, 156.4, 156.2, 154.6, 154.5, 149.3, 144.0, 143.9, 143.9, 140.8, 140.8, 127.7, 127.1, 125.2, 125.1, 123.0, 120.2, 65.6, 65.5, 49.2, 47.9, 47.1, 47.0, 46.8, 43.9, 43.7, 40.2, 38.0; *m*/*z* (TOF-HRMS-ES$^+$) 550.1605, $C_{26}H_{25}ClN_7O_5$ requires 550.1606.

(4) **{[2-(2-Amino-6-methoxyamino-purin-9-yl)-acetyl]-[2-(*9H*-flouren-9-ylmethoxy-carbonylamino)-ethyl]-amino}-acetic acid.** To a stirred solution of 3 (0.150 g, 0.27 mmol) in 1,4-dioxane (10 cm^3) under argon at 90°C was added methoxyamine hydrochloride (0.30 g, 3.6 mmol) and DIPEA (0.50 cm^3, 2.9 mmol). The mixture was stirred for 24 h and evaporated in vacuo. Purification by preparative HPLC (dissolved in 60:40 water:acetonitrile) or DCVC RP 18 silica gel (ø 3 x 5 cm) 100% water→100% acetonitrile, 10% increments/20 cm^3 fraction, yielding 0.087 g (57%) white solid (light yellow solid identical yield for RP18 silica gel). R_f 0.11 (40:60 methanol:acetonitrile); δ_H (300 MHz, DMSO-d_6) (13.19) 12.74 (1H, s, COO*H*), 10.70 (1H, br s, N*H*1), (8.22) 8.17 (1H, s, *H*8), 7.89 (2H, d, *J* 7.5 Hz, fluorenyl-*H*4 & 5), 7.68 (7.66) (2H, d, *J* 7 (6.5) Hz, fluorenyl-*H*1 & 8), 7.48-7.23 (5H, m, fluorenyl-*H*2,3,6,7 & Fmoc-N*H*), 6.99 (2H, br s, NH_2), 5.08 (4.91) (2H, s, CH_2N9), 4.40-4.17 (4.00) (5H, m, C*H*CH_2 & CH_2COOH), 3.81 (3H, s, CH_3), 3.53-3.04 (4H, m, CH_2CH_2); δ_C (75 MHz) (170.6) 170.2, 166.3, 156.4 (156.2), 153.4, 143.7 (143.7), 140.6 (140.5), 138.0, 127.4, 126.9, (124.9) 124.9, 119.9, 65.3 (65.2), 61.5, 47.7, 46.8, 46.6, 37.7, HRMS (ESI$^+$): 561.2184 (M + H^+ calc. 561.2205).

PNA oligomers were synthesized by standard methods and were purified by reversed phase HPLC. They were analysed for purity and identity by reversed phase HPLC and MALDI-TOF mass spectrometry, respectively. Tm measurements were performed in 100 mM NaCl, 0.1 mM EDTA, 10 mM Na_2HPO_4, pH 7.0 on a Cary 300Bio spectrophotometer at a heating rate of 0.5°C/min.

Acknowledgements

This work was supported by The Danish Research Agency (Center for Organic Chemical Biotechnology) and the European commission (contract no QLK3-CT-2000-00634). Ms. Jolanta Ludvigsen and Ms. Annette Jørgensen are thanked for expert technical assistance.

References

1. Lohse J, Dahl O, Nielsen PE. Double duplex invasion by peptide nucleic acid: a general principle for sequence-specific targeting of double-stranded DNA. Proc Natl Acad Sci USA 1999; 96:11804-8.
2. Lonkar P, Kim KH, Kuan JY, Chin JY, Rogers FA, Knauert MP, et al. Targeted correction of a thalassemia-associated beta-globin mutation induced by pseudocomplementary peptide nucleic acids. Nucleic Acids Res 2009; 37:3635-44.
3. Kim KH, Nielsen PE, Glazer PM. Site-directed gene mutation at mixed sequence targets by psoralen-conjugated pseudo-complementary peptide nucleic acids. Nucleic Acids Res 2007; 35:7604-13.
4. Izvolsky KI, Demidov VV, Nielsen PE, Frank-Kamenetskii MD. Sequence-specific protection of duplex DNA against restriction and methylation enzymes by pseudocomplementary PNAs. Biochemistry 2000; 39:10908-13.
5. Demidov VV, Protozanova E, Izvolsky KI, Price C, Nielsen PE, Frank-Kamenetskii MD. Kinetics and mechanism of the DNA double helix invasion by pseudocomplementary peptide nucleic acids. Proc Natl Acad Sci USA 2002; 99:5953-8.
6. Protozanova E, Demidov VV, Nielsen PE, Frank-Kamenetskii MD. Pseudocomplementary PNAs as selective modifiers of protein activity on duplex DNA: the case of type IIs restriction enzymes. Nucleic Acids Res 2003; 31:3929-35.

7. Katada H, Chen HJ, Shigi N, Komiyama M. Homologous recombination in human cells using artificial restriction DNA cutter. Chem Commun 2009; 6545-7.
8. Ito K, Katada H, Shigi N, Komiyama M. Site-selective scission of human genome by artificial restriction DNA cutter. Chem Commun 2009; 6542-4.
9. Katada H, Komiyama M. Artificial restriction DNA cutters as new tools for gene manipulation. Chembiochem 2009; 10:1279-88.
10. Miyajima Y, Ishizuka T, Yamamoto Y, Sumaoka J, Komiyama M. Origin of high fidelity in target-sequence recognition by PNA-Ce(IV)/EDTA combinations as site-selective DNA cutters. J Am Chem Soc 2009; 131:2657-62.
11. Gamper HB Jr, Arar K, Gewirtz A, Hou YM. Unrestricted accessibility of short oligonucleotides to RNA. RNA 2005; 11:1441-7.
12. Lahoud G, Timoshchuk V, Lebedev A, Arar K, Hou YM, Gamper H. Properties of pseudo-complementary DNA substituted with weakly pairing analogs of guanine or cytosine. Nucleic Acids Res 2008; 36:6999-7008.
13. Lahoud G, Timoshchuk V, Lebedev A, de Vega M, Salas M, Arar K, et al. Enzymatic synthesis of structure-free DNA with pseudo-complementary properties. Nucleic Acids Res 2008; 36:3409-19.
14. Brown DM, Lin PKT. Synthesis and duplex stability of oligonucleotides containing adenine-guanine analogues. Carbohydr Res 1991; 216:129-39.
15. Christensen L, Hansen HF, Koch T, Nielsen PE. Inhibition of PNA Triplex Formation by N4-Benzoylated Cytosine. Nucleic Acids Res 1998; 26:2735-9.
16. Dueholm KL, Egholm M, Behrens C, Christensen L, Hansen HF, Vulpius T, et al. Synthesis of Peptide Nucleic Acid Monomers Containing the Four Natural Nucleobases: Thymine, Cytosine, Adenine and Guanine and Their Oligomerization. J Org Chem 1994; 59:5767-73.
17. Thomson SA, Josey JA, Cadilla R, Gaul MD, Hassman CF, Luzzio MJ, et al. Fmoc Mediated Synthesis of Peptide Nucleic Acids. Tetrahedron 1995; 51:6179-94.
18. Jones LW, Major RT. Substituted O-Alkyl Hydroxylamines Chemically Related to Medicinally Valuable Amines. J Am Chem Soc 1927; 49:1527-40.
19. Pedersen DS, Rosenbohm C. Dry Column Vacuum Chromatography. Synthesis-Stuttgart 2001; 2431-4.